Supporting Information

Chromatin unfolding by epigenetic modifications explained by dramatic impairment of internucleosome interactions: a multiscale computational study

Chromatin coarse-grained model

Our work includes Monte Carlo simulations of 24-nucleosome arrays carried out with a our coarse-grained chromatin model[1-12]. The model has been described in detail in[3,7,10], and below we summarize the strategies used to treat each oligonucleosome component:

Nucleosome cores. The nucleosome protein core, excluding histone tails, with wrapped DNA is modelled as a rigid irregular body with 300 Debye-Hückel charges uniformly distributed on the nucleosome molecular surface. The charges are optimized to reproduce the full atom electric field around the nucleosome core by the Discrete Surface Charge Optimization (DiSCO) algorithm[2], which solves the complete nonlinear Poisson-Boltzmann equation.

Flexible histone tails. Our original model considers the ten histone tails protruding out of each core (the N-termini of each H2A, H2B, H3, and H4, plus the C-termini of each H2A) as flexible chains of beads with the first bead rigidly attached to the parent core. Each bead comprises 5 amino acids and its centred at the C_β atom of the middle amino acid. Each tail chain is assigned a customized intramolecular force field comprising bond stretching and bond-angle bending terms[1,3]. The parameters for this force field (i.e., equilibrium bond lengths and bond angles and the related force constants) are optimized to reproduce the configurational properties of the atomistic histone tails[1,3]. The charges of the tail beads are also optimized to reproduce the atomistic properties of the amino acids they represent. That is, each bead is assigned a charge equal to the sum of the charges on its five amino acids, multiplied by a scaling factor close to unity (1.12 for 0.15M NaCl used here) that accounts for salt dependence in the effective charge.

Folded histone tails. We assign one bead per each 5 amino acids and centre it at the C_β atom of the middle amino acid using as reference structure the centroid of the highest populated folded cluster obtained in our REMD simulations. We limit tail flexibility by increasing the stretching, bending and torsional inter-tail-bead force constants by a factor of 100. The tails can spontaneously fold/unfold through our tailored MC move (see Supporting Material) that attempts transition between folded and flexible tails.

DNA linkers. The DNA that connects consecutive nucleosomes is treated as a chain of spherical beads that have a salt-dependent charge parameterized using the Stigter procedure[13]. The mechanical properties of the linker DNA chains are also considered, and described with the combined wormlike-chain (WLC) model[14] of Jian et al. [15,16]. The equilibrium DNA inter-bead segment is 3 nm or 9 bp, thus to model NRLs of 182 bp and 209-bp we use 3 and 6 DNA beads (4 and 7 segments) per linker, respectively. The exiting and entering DNA linkers attached to the nucleosome define an angle of 108°, which corresponds to the 147 DNA base pairs tightly wound ~1.7 times around the core[10].

Solvent and ionic environment. The water around the oligonucleosome is treated implicitly as a continuum. The screening of electrostatic interactions due to the presence of monovalent ions in solution (0.15 M NaCl) is treated using a Debye-Hückel potential (electrostatic screening length of 1.27 nm^{-1}) [3] and, as described above, with the charges on each component parameterized considering salt-dependent screening.

To prevent overlap among chromatin components, each nucleosome charge, linker DNA bead, and histone tail bead are assigned an excluded volume. Specific expressions for the oligonuclesome energy and all values of parameters can be found in [3,7,10].

Monte Carlo algorithm for the simulation of chromatin

We sample our 24-nucleosome chromatin conformations at constant temperature using a Monte Carlo (MC) procedure with five different moves.

The first three are our standard global pivot, local translation, and local rotation moves, which focus on the conformational sampling of the main oligonucleosome chain (nucleosomes joined by DNA beads). The global pivot move is implemented by randomly choosing one linker DNA bead or nucleosome core and a random axis passing through the chosen component. The shorter part of the oligonucleosome about this axis is rotated by an angle chosen from a uniform distribution within a range set so that the acceptance probability is ~35%. The local translation and rotation moves also select randomly an oligonucleosome chain component (linker DNA bead or core) and an axis passing through it. In the translation/rotation move, the component is then moved/rotated along/about the axis by a distance/angle sampled from a uniform distribution also chosen so that the acceptance probability is ~35%.

The fourth is our new tail folding/unfolding move, which implements transitions between folded and unfolded tails. This move randomly selects a histone tail chain, and either folds it and rigidifies it, or unfolds it and allows it to become flexible with probabilities P and 1-P, respectively. By changing the value of P, we control the concentration of folded and unfolded tails. The different chromatin conformations in the resulting equilibrium ensemble have a fixed concentration of folded/unfolded tails; however, the specific locations of the folded/unfolded tails change among the different conformations. The resulting equilibrium ensemble thus mimics an array of chromatin fibers in which the tails transiently fold and unfold.

The fifth is our tail regrowth move, which is implemented to sample flexible histone-tail conformations based on the configurational bias MC method [17]. This move randomly selects a histone tail chain defined as a flexible tail, and regrows it bead-by-bead using the Rosenbluth scheme [18]. To prevent histone tail beads from penetrating the nucleosome core, the volume enclosed within the nucleosome surface is discretized, and any trial configurations that place the beads within this volume are rejected automatically.

The first three moves are accepted or rejected based on the Metropolis criterion. The pivot, translation, rotation, tail folding/unfolding, and tail regrowth moves are attempted with probabilities of 0.2, 0.1, 0.1, 0.2, and 0.4, respectively.

Calculation of the absolute and relative packing ratios

The absolute packing ratio is a measurement of oligonucleosomes compactness, and is the defined as the number of nucleosomes per 11nm of oligonucleosome length. To calculate this packing ratio, we compute the length of the oligonucleosome fiber axis passing. We define the fiber axis as a 3-dimensional parametric curve

$$\mathbf{r}^{ax}(i) = \left(r_1^{ax}(i), r_2^{ax}(i), r_3^{ax}(i)\right) \qquad (2)$$

where $r_j^{ax}(i)$ are three functions that map the center positions of the i^{th} nucleosome $\mathbf{r}_i = \left(r_{i,1}, r_{i,2}, r_{i,3}\right)$. We approximate these functions with second order polynomials of the form

$$r_j^{ax}(i) \approx P_j(i) = \sum_{k=1}^{3} p_{k,j} i^{(3-k)} \qquad (3)$$

by fitting the data sets $[\boldsymbol{r}_i]$ by a least-squares procedure. We determine the coefficients of the polynomial $P_j(i)$ by minimizing the sum of the squares of the residuals

$$l_j = \sum_{i=1}^{N_C} \left(r_{i,j} - P_j(i)\right)^2 \qquad (4)$$

where N_C gives the number of nucleosome cores in an oligonucleosomes. This residual function accounts for the differences between a proposed polynomial fit and the observed

nucleosome positions. After determining the polynomial coefficients, we use Eq. (3) to produce N_C points per spatial dimension and compute the fiber length L_{fiber} as follows:

$$L_{\text{fiber}} = \sum_{i=1}^{(N_C-1)/2} |\mathbf{r}^{ax}(2i-1) - \mathbf{r}^{ax}(2i+1)| \quad (5)$$

where the distances are between every two consecutive nucleosome centres. The absolute packing ratio P_{A} is then calculated as the number of cores multiplied by 11nm/L_{fiber}. In addition, we report relative packing ratios to describe the loss of compaction upon histone tail folding more easily. We have defined these relative packing ratios as

$$P_{\text{R}} = \left(\frac{P_{\text{A}} - P_{\text{O}}}{P_{\text{C}} - P_{\text{O}}}\right) \times 100\% \quad (6)$$

where P_{O} is the absolute packing ratio calculated for an open oligonucleosome modelled at low monovalent salt (0.01 M NaCl), no LHs, and 100% flexible histone tails; and P_{C} is the absolute packing ratio calculated for a fully condensed oligonucleosome modelled at high monovalent salt (0.15 M NaCl), no LHs, and 100% flexible histone tails. Fully compact fibers give a relative packing ratio of 100%, while the low salt open fibers produce packing ratios of 0%.

Frequency of tail-mediated interactions.

We measure the fraction of configurations that tails of a specific kind t (t = H4, H3, H2B, and H2A) in a chromatin chain are 'in contact' with a specific component c of the chromatin chain (c = a non-parental nucleosome or a non-parental DNA linkers) (Fig. S5b). To do this, we construct two-dimensional matrices with the following elements

$$T'(t,c) = mean\left[\frac{1}{N_C N}\sum_{i\in I_C}\sum_{j=1}^{N}\partial_{i,j}^{t,c}(M)\right]. \quad (7)$$

Here N_C is the number of nucleosomes in the chromatin array, N the total number of chromatin components (nucleosomes and linker DNAs), and I_C indicates a nucleosome particle within the chromatin chain. The mean above is taken over the converged MC configurations used for statistical analysis and

$$\partial_{i,j}^{t,c}(M) = \begin{cases} 1 & \text{if } j \text{ is a c-type component 'in contact' with} \\ & \text{a tail of kind t of nucleosome i at MC frame } M \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

For a MC frame M, we consider a specific t-kind tail of core i to be either free or in contact with only one of the N chromatin components of the oligonucleosome chain. The t-tail is in contact with a component of type c if the shortest distance between its beads and the beads or core charges of c is smaller than the shortest distance to any other type of component and also smaller than the relevant tail-component excluded volume distance (see parameters in [10]). The resulting normalized patterns of interactions provide crucial information into the frequency by which different tails mediate chromatin interactions.

SUPPORTING TABLES

System no.	System	Protocol	Force field	Water model	Simulation length
1	H4 tail	REMD	AMBER99SB*-ILDN	TIP3P	56 replicas x 500 ns each
2		REMD	AMBER99SB	TIP3P	56 replicas x 500 ns each
3		REMD	CHARMM36	TIP3P	56 replicas x 500 ns each
4	H4 K16Ac tail	REMD	AMBER99SB*-ILDN + Papageorgiou's KAc parameters	TIP3P	56 replicas x 500 ns each
5		REMD	AMBER99SB + Papageorgiou's KAc parameters	TIP3P	56 replicas x 500 ns each
6		REMD	CHARMM36 + Dejaegere's KAc parameters	TIP3P	56 replicas x 500 ns each
7	H4 K12Ac tail	REMD	AMBER99SB*-ILDN + Papageorgiou's KAc parameters	TIP3P	56 replicas x 500 ns each
8	H4 K12,16Ac tail	REMD	AMBER99SB*-ILDN + Papageorgiou's KAc parameters	TIP3P	56 replicas x 500 ns each
9	H4 K5,8,12Ac tail	REMD	AMBER99SB*-ILDN + Papageorgiou's KAc parameters	TIP3P	64 replicas x 500 ns each
10	H4 K5,8,12,16Ac tail	REMD	AMBER99SB*-ILDN + Papageorgiou's KAc parameters	TIP3P	64 replicas x 500 ns each
11	H3 tail	REMD	AMBER99SB*-ILDN	TIP3P	56 replicas x 500 ns each
12	H3 tail	REMD	AMBER99SB	TIP3P	56 replicas x 500 ns each
13	H3 K14Ac tail	REMD	AMBER99SB*-ILDN + Papageorgiou's KAc parameters	TIP3P	56 replicas x 500 ns each

14	H2B tail	REMD	AMBER99SB*-ILDN	TIP3P	56 replicas x 500 ns each
15	H2B tail	REMD	AMBER99SB	TIP3P	56 replicas x 500 ns each
16	H2B K20Ac tail	REMD	AMBER99SB*-ILDN + Papageorgiou's KAc parameters	TIP3P	56 replicas x 500 ns each
17	H2B K5,12,15,20Ac tail	REMD	AMBER99SB*-ILDN + Papageorgiou's KAc parameters	TIP3P	56 replicas x 500 ns each
18	H2A tail	REMD	AMBER99SB*-ILDN	TIP3P	56 replicas x 500 ns each
19	H2A tail	REMD	AMBER99SB	TIP3P	56 replicas x 500 ns each
20	H2AC tail	REMD	AMBER99SB*-ILDN	TIP3P	56 replicas x 500 ns each
21	H2AC tail	REMD	AMBER99SB	TIP3P	56 replicas x 500 ns each
22	H4 tail	Chemical shift restraints	AMBER99SB*-ILDN	TIP3P	8 replicas x 500 ns each
23	H4 tail	Chemical shift restraints	CHARMM36	TIP3P	8 replicas x 500 ns each
24	H4 tail	Chemical shift restraints + MetaDynamics	AMBER99SB*-ILDN	TIP3P	8 replicas x 500 ns each
25	H4 tail	Free MD	AMBER99SB*-ILDN	TIP3P	1 μs
26	H4 K16Ac tail	Free MD	AMBER99SB*-ILDN + Papageorgiou's KAc parameters	TIP3P	1 μs
27	H3 tail	Free MD	AMBER99SB*-ILDN	TIP3P	1 μs
28	H2B tail	Free MD	AMBER99SB*-ILDN	TIP3P	1 μs
29	H2A tail	Free MD	AMBER99SB*-ILDN	TIP3P	1 μs
30	H2AC tail	Free MD	AMBER99SB*-ILDN	TIP3P	1 μs
31	Dinucleosome	Free MD with	AMBER99SB*-	TIP3P	4 μs

	with full wild-type tails	virtual sites	ILDN + AMBER99+parmBSC0		
32	Dinucleosome with H4 K16Ac tail, H3 K14Ac tail, and wild type H2B, H2A and H2AC tails	Free MD with virtual sites	AMBER99SB*-ILDN + Papageorgiou's KAc parameters + AMBER99+parmBSC0	TIP3P	4 μs

Table S1. List of explicit solvent all-atom molecular dynamics simulations performed in this work.

System no.	NRL	Folded tail concentration / other info	Salt Concentration	Sampling
1	182 bp	0%	0.01M	12 trajectories x 50 million MC steps
2	182 bp	0%	0.15M	12 trajectories x 50 million MC steps
3	182 bp	5% all tails	0.15M	12 trajectories x 50 million MC steps
4	182 bp	10% all tails	0.15M	12 trajectories x 50 million MC steps
5	182 bp	25% all tails	0.15M	12 trajectories x 50 million MC steps
6	182 bp	50% all tails	0.15M	12 trajectories x 50 million MC steps
7	182 bp	75% all tails	0.15M	12 trajectories x 50 million MC steps
8	182 bp	90% all tails	0.15M	12 trajectories x 50 million MC steps
9	182 bp	100% all tails	0.15M	12 trajectories x 50 million MC steps
10	182 bp	5% H4	0.15M	12 trajectories x 50 million MC steps
11	182 bp	10% H4	0.15M	12 trajectories x 50 million MC steps
12	182 bp	25% H4	0.15M	12 trajectories x 50 million MC steps
13	182 bp	50% H4	0.15M	12 trajectories x 50 million MC steps

14	182 bp	75% H4	0.15M	12 trajectories x 50 million 15MC steps
15	182 bp	90% H4	0.15M	1216 trajectories x 50 million MC steps
16	182 bp	100% H4	0.15M	12 trajectories x 50 million MC steps
17	182 bp	5% H3	0.15M	12 trajectories x 50 million MC steps
18	182 bp	10% H3	0.15M	12 trajectories x 50 million MC steps
19	182 bp	25% H3	0.15M	12 trajectories x 50 million MC steps
20	182 bp	50% H3	0.15M	12 trajectories x 50 million MC steps
21	182 bp	75% H3	0.15M	12 trajectories x 50 million MC steps
22	182 bp	90% H3	0.15M	12 trajectories x 50 million MC steps
23	182 bp	100% H3	0.15M	12 trajectories x 50 million MC steps
24	182 bp	5% H2B	0.15M	12 trajectories x 50 million MC steps
25	182 bp	10% H2B	0.15M	12 trajectories x 50 million MC steps
26	182 bp	25% H2B	0.15M	12 trajectories x 50 million MC steps
27	182 bp	50% H2B	0.15M	12 trajectories x 50 million MC steps
28	182 bp	75% H2B	0.15M	12 trajectories x 50 million MC steps
29	182 bp	90% H2B	0.15M	12 trajectories x 50 million MC steps
30	182 bp	100% H2B	0.15M	12 trajectories x 50 million MC steps
31	182 bp	5% H2A	0.15M	12 trajectories x 50 million MC steps
32	182 bp	10% H2A	0.15M	12 trajectories x 50 million

				MC steps
33	182 bp	25% H2A	0.15M	12 trajectories x 50 million MC steps
34	182 bp	50% H2A	0.15M	12 trajectories x 50 million MC steps
35	182 bp	75% H2A	0.15M	12 trajectories x 50 million MC steps
36	182 bp	90% H2A	0.15M	12 trajectories x 50 million MC steps
37	182 bp	100% H2A	0.15M	12 trajectories x 50 million MC steps
38	182 bp	0% / charge of H4K16Ac bead reduced by 1e	0.15M	12 trajectories x 50 million MC steps
39	182 bp	0% / charge of H3K14Ac bead reduced by 1e	0.15M	12 trajectories x 50 million MC steps
40	209 bp	0%	0.01M	12 trajectories x 50 million MC steps
41	209 bp	0%	0.15M	12 trajectories x 50 million MC steps
42	209 bp	5% all tails	0.15M	12 trajectories x 50 million MC steps
43	209 bp	10% all tails	0.15M	12 trajectories x 50 million MC steps
44	209 bp	25% all tails	0.15M	12 trajectories x 50 million MC steps
45	209 bp	50% all tails	0.15M	12 trajectories x 50 million MC steps
46	209 bp	75% all tails	0.15M	12 trajectories x 50 million MC steps
47	209 bp	90% all tails	0.15M	12 trajectories x 50 million MC steps
48	209 bp	100% all tails	0.15M	12 trajectories x 50 million MC steps
49	209 bp	5% H4	0.15M	12 trajectories x 50 million MC steps
50	209 bp	10% H4	0.15M	12 trajectories

				x 50 million MC steps
51	209 bp	25% H4	0.15M	12 trajectories x 50 million MC steps
52	209 bp	50% H4	0.15M	12 trajectories x 50 million MC steps
53	209 bp	75% H4	0.15M	12 trajectories x 50 million MC steps
54	209 bp	90% H4	0.15M	12 trajectories x 50 million MC steps
55	209 bp	100% H4	0.15M	12 trajectories x 50 million MC steps
56	209 bp	5% H3	0.15M	12 trajectories x 50 million MC steps
57	209 bp	10% H3	0.15M	12 trajectories x 50 million MC steps
58	209 bp	25% H3	0.15M	12 trajectories x 50 million MC steps
59	209 bp	50% H3	0.15M	12 trajectories x 50 million MC steps
60	209 bp	75% H3	0.15M	12 trajectories x 50 million MC steps
61	209 bp	90% H3	0.15M	12 trajectories x 50 million MC steps
62	209 bp	100% H3	0.15M	12 trajectories x 50 million MC steps
63	209 bp	5% H2B	0.15M	12 trajectories x 50 million MC steps
64	209 bp	10% H2B	0.15M	12 trajectories x 50 million MC steps
65	209 bp	25% H2B	0.15M	12 trajectories x 50 million MC steps
66	209 bp	50% H2B	0.15M	12 trajectories x 50 million MC steps
67	209 bp	75% H2B	0.15M	12 trajectories x 50 million MC steps
68	209 bp	90% H2B	0.15M	12 trajectories x 50 million MC steps
69	209 bp	100% H2B	0.15M	12 trajectories

				x 50 million MC steps
70	209 bp	5% H2A	0.15M	12 trajectories x 50 million MC steps
71	209 bp	10% H2A	0.15M	12 trajectories x 50 million MC steps
72	209 bp	25% H2A	0.15M	12 trajectories x 50 million MC steps
73	209 bp	50% H2A	0.15M	12 trajectories x 50 million MC steps
74	209 bp	75% H2A	0.15M	12 trajectories x 50 million MC steps
75	209 bp	90% H2A	0.15M	12 trajectories x 50 million MC steps
76	209 bp	100% H2A	0.15M	12 trajectories x 50 million MC steps
77	209 bp	0% / charge of H4K16Ac bead reduced by 1e	0.15M	12 trajectories x 50 million MC steps
78	209 bp	0% / charge of H3K14Ac bead reduced by 1e	0.15M	12 trajectories x 50 million MC steps
79	191 bp	0%	0.15M	12 trajectories x 50 million MC steps
80	200 bp	0%	0.15M	12 trajectories x 50 million MC steps

Table S2. List of coarse-grained 24-nucleosome arrays without linker histones simulated in this work.

Tail	**Number of amino acids (N)**	**Total % SS**	**Persistence length (L_p)**	**Contour length (L=N*0.38 nm)**
H4 WT	26	8.53±0.76	0.44 (±0.02) nm	9.88 nm
H3 WT	38	14.15±1.94	0.79 (±0.02) nm	14.44 nm
H2B WT	23	13.85±3.76	0.69 (±0.02) nm	8.74 nm
H2A WT	14	4.71±0.11	0.76 (±0.02) nm	5.32 nm

H2AC WT	9	7.63±0.08	0.60 (±0.01) nm	3.42 nm
Titin PEVK11 peptide (exp)	11	---	0.63 (±0.01) nm	4.18 nm
Titin PEVK21 peptide (exp)	21	---	0.48 (±0.02) nm	7.98 nm
Polyproline (exp) [19]	6,9,11,12,13,15,20,23,27,33,40	---	4.4 (±0.9) nm	2.28-15.20 nm

Table S3: Persistence and contour length of histone tails.

Protein	**Total % SS**	**Lp**	**Lp increase**
H4 WT	**8.53±0.76**	0.44 (±0.02) nm	**--**
H4 K16Ac	12.10±0.97	0.62 (±0.02) nm	41%
H4 K12Ac	7.36±0.70	0.60 (±0.01) nm	36%
H4 diAc	10.33±0.85	0.58 (±0.02) nm	32%
H4 triAc	8.87±0.88	0.57 (±0.01) nm	30%
H4 tetraAc	8.61±0.80	0.57 (±0.02) nm	30%
H3 WT	**14.15±1.94**	0.79 (±0.02) nm	**--**
H3 K14Ac	20.49±2.64	0.89 (±0.02) nm	13%
H2B WT	**13.85±3.76**	0.69 (±0.02) nm	**--**
H2B K20Ac	13.60±1.05	0.74 (±0.01) nm	7%
H2B tetraAc	18.10±2.19	0.98 (±0.09) nm	42%

Table S4. Persistence-to-contour-length values for different lysine-acetylated histone tails.

SUPPORTING FIGURES

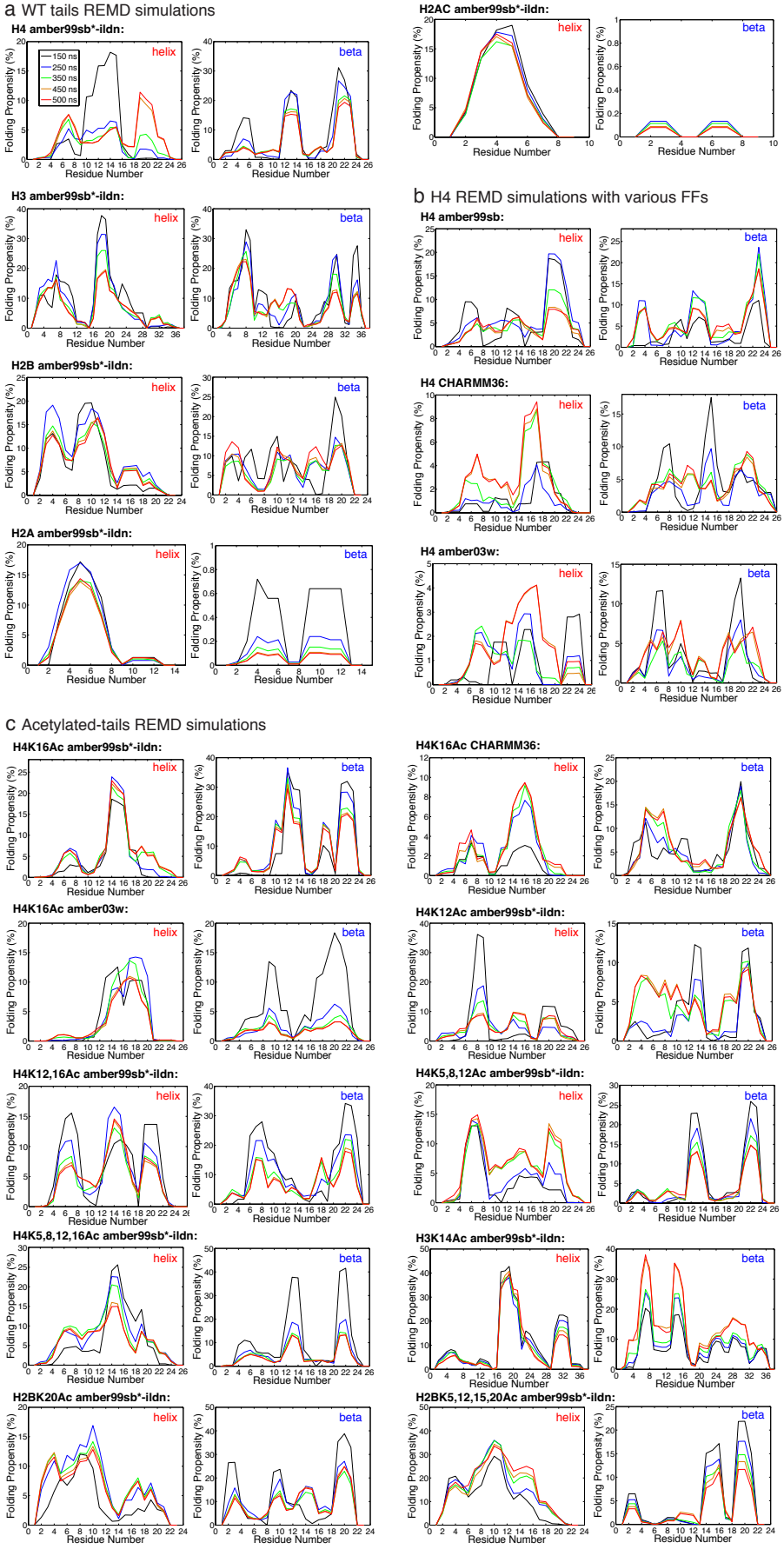

Figure S1. Assessment of the convergence of the REMD simulations. The assessment was made by monitoring the changes in the α-helical (columns 1 and 3) and **β**-strand (columns 2 and 4) folding propensity patterns for the lowest temperature replica over simulation time. The first 100 ns were discarded for equilibration, and the percentages of folded conformations per residue (folding propensity) computed over 100-to-150 ns (labeled 150 ns in black), 100-to-250 ns (labeled 250 ns in blue), 100-to-350 ns (labeled 350 ns in green), 100-to-450 ns (labeled 450 ns in orange), and 100-to-500 ns (labeled 500 ns in red) are shown. Plots are for the: (a) WT histone tails, (b) the H4 tail with different force fields, and (c) the acetylated histone tails.

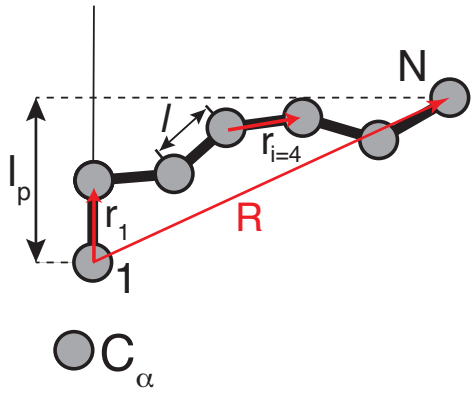

Figure S2. Schematic illustration of the persistence length calculation. See equation (1).

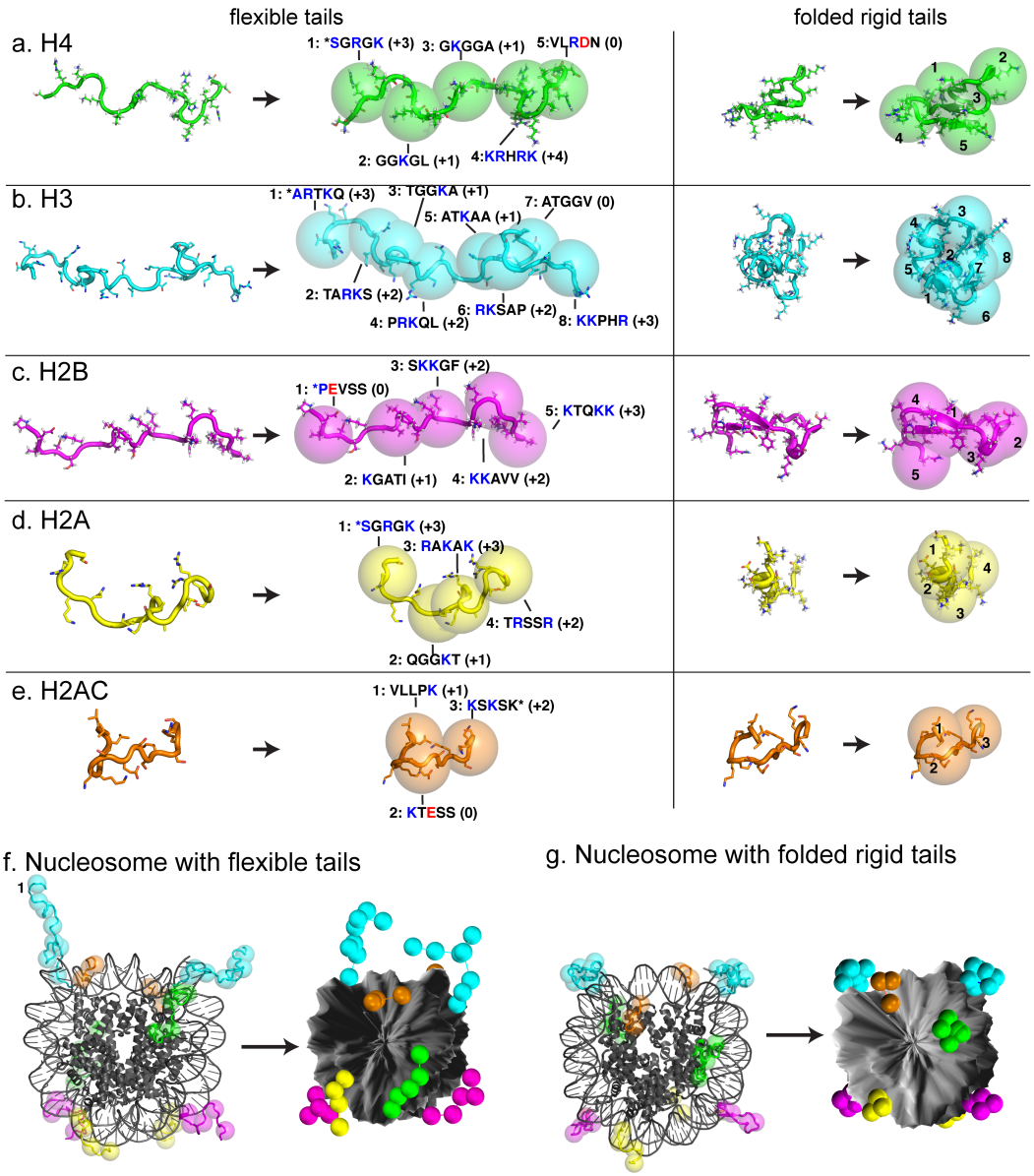

Figure S3. Correspondence between coarse-grained histone tail models and all-atom structures. (a-e) For each histone tail, the figure presents all-atom models and overlays of the locations of the histone tail beads on top of the all-atom models for: unstructured histone tails (left) and the most populated structured arrangement obtained in our REMD simulations (right). Histone tails are colour green (H4), cyan (H3), magenta (H2B), yellow (H2A), and orange (H2A C-tail). Each bead represents five consecutive amino acids and is centred on the beta carbon of the middle amino acid. Each bead of the flexible tail models has been labelled with its bead number (numbering started from the N-terminus), the sequence of amino acids represented by each bead (neutral amino acids are written in black, positively charged ones in blue, and negatively charged ones in red), and the total charge of the bead. Here, the asterisk indicates that the charge of the N- and C-termini has been considered. (f-g) Attachment of the flexible and folded histone tail model into an all-atom nucleosome and its corresponding coarse grained representation. Histone cores and nucleosomal DNA are depicted in grey.

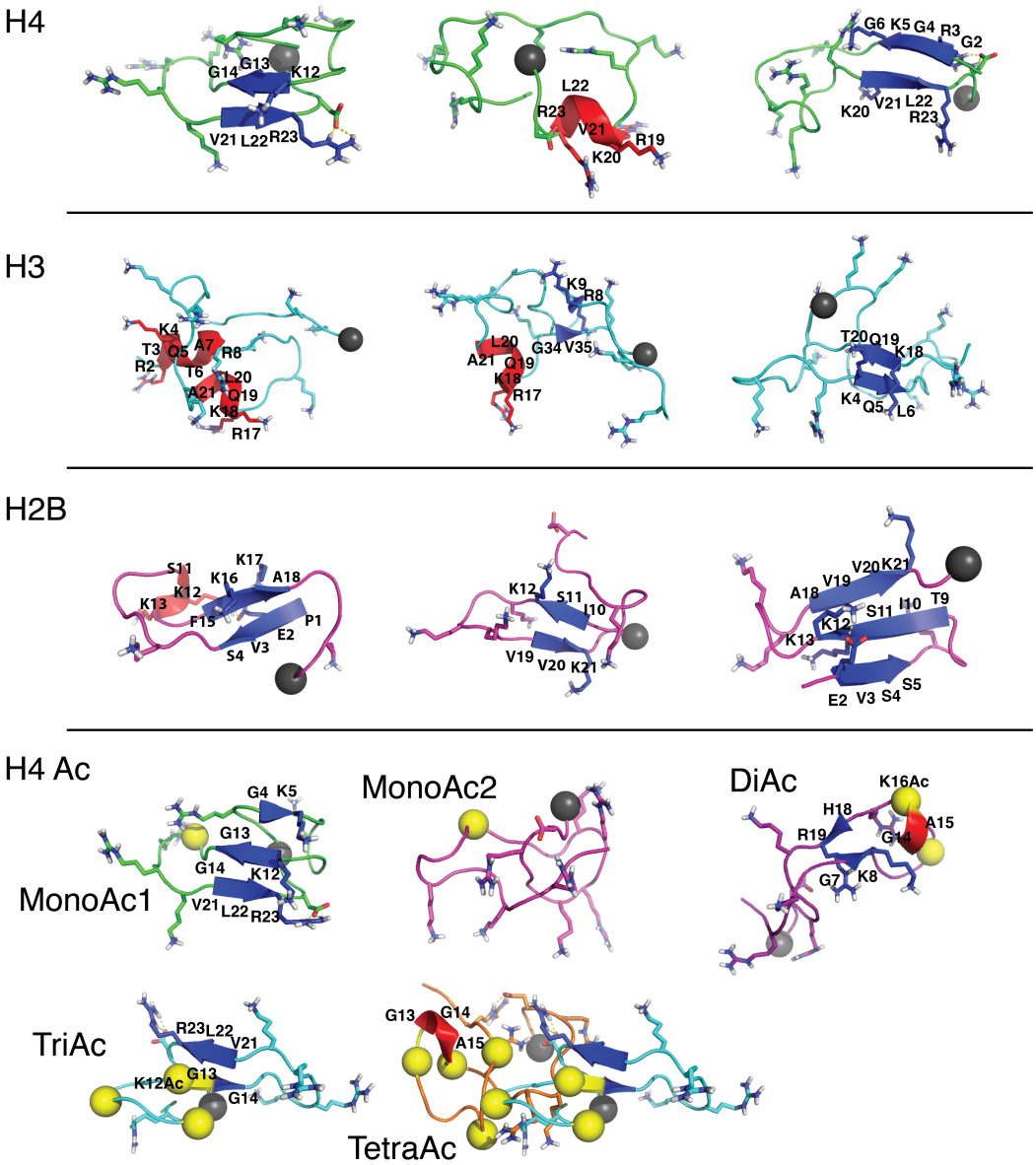

Figure S4. Histone tails' most common folded structures. The top three panels show the structures of the three most populated clusters for the H4 (green), H3 (cyan), and H2B tails (magenta). The bottom panel shows the structure of the most populated cluster for the following H4 lysine acetylated versions: K16Ac (MonoAc1; green), K12Ac (MonoAc2; magenta), K12,16Ac (DiAc; purple), K5,8,12Ac (TriAc; cyan) and K5,8,12,16Ac (TetraAc; orange). The alpha helical motifs are highlighted in red and the beta motifs in blue. The residues involved in secondary structural motifs are labelled and the side chains are drawn with sticks with the polar hydrogens removed for clarity. The last residue is indicated with a black sphere and the acetylated lysines with a yellow sphere.

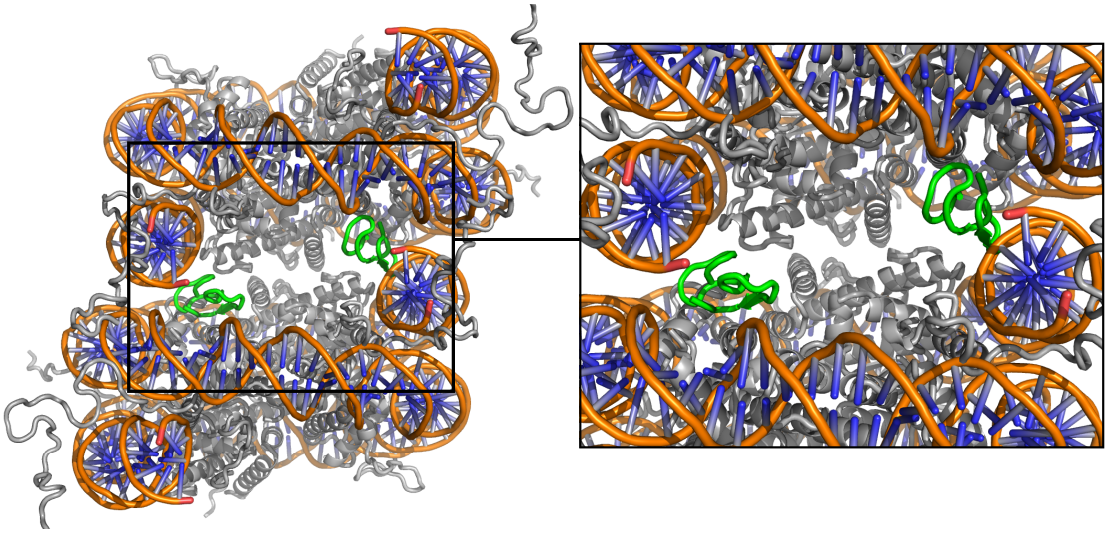

Figure S5. Folded H4 tails within a dinucleosome. This figure shows how the common H4 folded structures would fit within two closely interacting nucleosomes. We have constructed this model by placing two 1KX5 nucleosomes on top of each other using the geometry of stacked nucleosomes in the tetranucleosome crystal structure, and replacing the H4 tails with the most populated structures found in our REMD simulations.

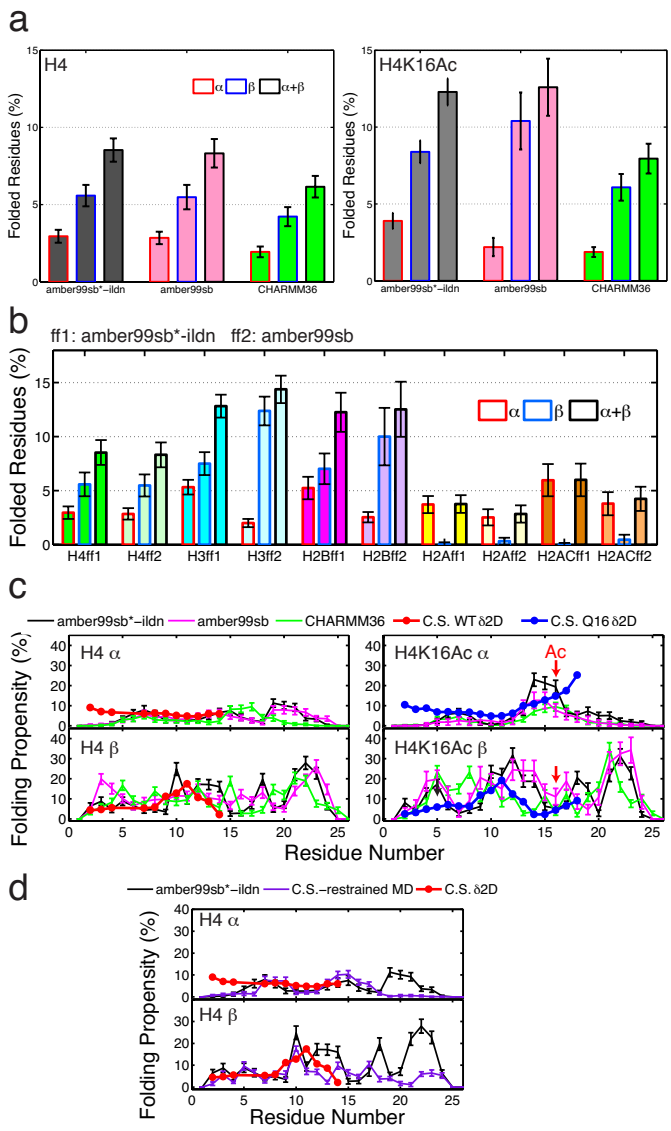

Figure S6. Analysis of the effects of the force field on the results of the simulations. (a) Ensemble average and standard deviation of the percentage of residues of H4 and H4K16Ac that adopt secondary structural elements assessed from the lowest temperature replica in our REMD simulations using different force fields. For H4 and H4 K16Ac, we used three of the latest force fields for proteins: (1) AMBER99SB*-ILDN, (2) AMBER99SB, and (3) CHARMM36. Lysine acetylated parameters taken [20] for AMBER99SB*-ILDN and AMBER99SB, and from [21] for CHARMM36. (b) Ensemble average and standard deviation of the percentage of residues of all wild-type tails to adopt secondary structural elements assessed from the lowest temperature replica in our REMD simulations using two different force fields: (1) AMBER99SB*-ILDN, and (2) AMBER99SB. Figures (a) and (b) show that the important trend of increased secondary structure, specially β motifs, upon acetylation remains for all force fields analysed. (c) Folding propensity per residue for H4 and H4 K16Ac and the three force fields used in (a) compared with the folding propensities calculated using experimental chemical shifts (red and blue, δ2D method[22]). To be consistent with the experiment, for this comparison we classify the α structures as those containing either α, 3_{10}, or π helices, and the β structures as those containing either isolated β bridges or extended conformations. Note that we compare the H4K16Ac folding propensities with the experimental H4K16Q mutation instead; however, how well the K16Q mutation mimics K16Ac is controversial, because while the acetylated version opens chromatin, the mutation does not alter chromatin compaction [23]. (d) Secondary structure motifs obtained from MD simulations and predicted based on experimental chemical shifts. AMBER99SB*-ILDN force field (AMB[24]): from 500 ns REMD simulations; C.S-restrained MD: from eight 500 ns replicas and

metadynamics on the end-to-end distance and number of hydrogen bonds[25]; C.S. δ2D: predicted from experimentally determined chemical shifts of the H4 tail in a nucleosome in solution using the δ2D predictor.

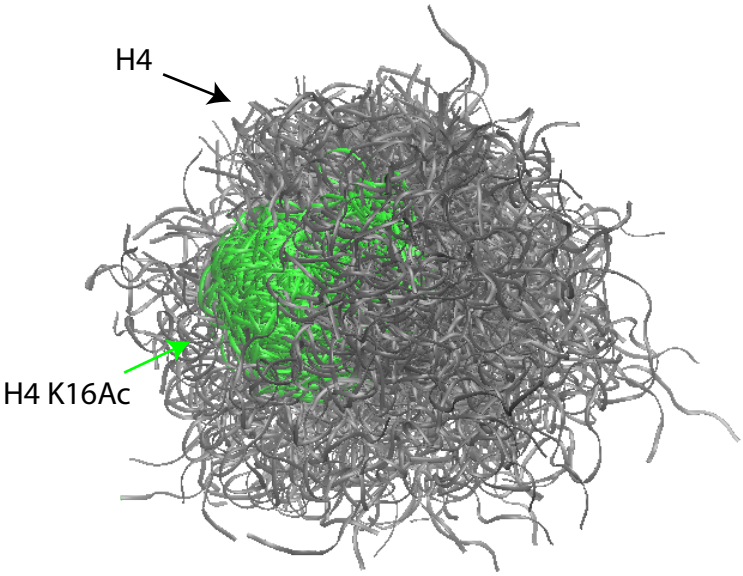

Figure S7. Spatial distribution of H4 and H4 K16Ac tails during a 1 μs-long MD trajectory. The last amino acid of all frames were aligned together. The H4 and H4K16Ac tails are shown as grey and green ribbons, respectively.

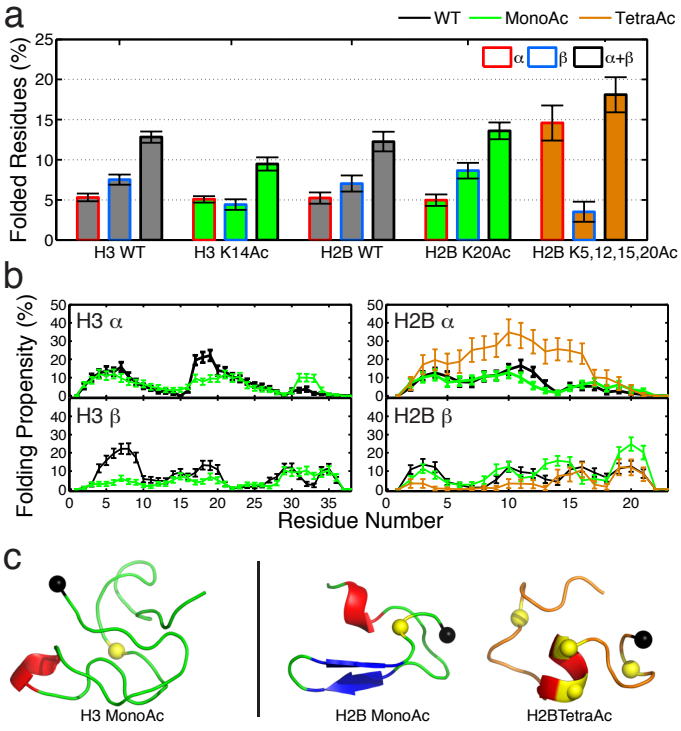

Figure S8. Effect of the acetylation of different lysine residues in the H3 and H2B tails. (a) Percentage of residues in various lysine-acetylated tails with secondary structure motifs. (b) Effect of acetylation in the folding propensity for each residue separated by α-helical and beta strand structural motifs. (c) Illustration of highest populated clusters with folded resides. α-helical motifs are coloured in red, while beta conformations in blue. The black sphere indicates the last residue of the N-tail (point of attachment to the nucleosome), while the yellow sphere denotes the acetylated lysine.

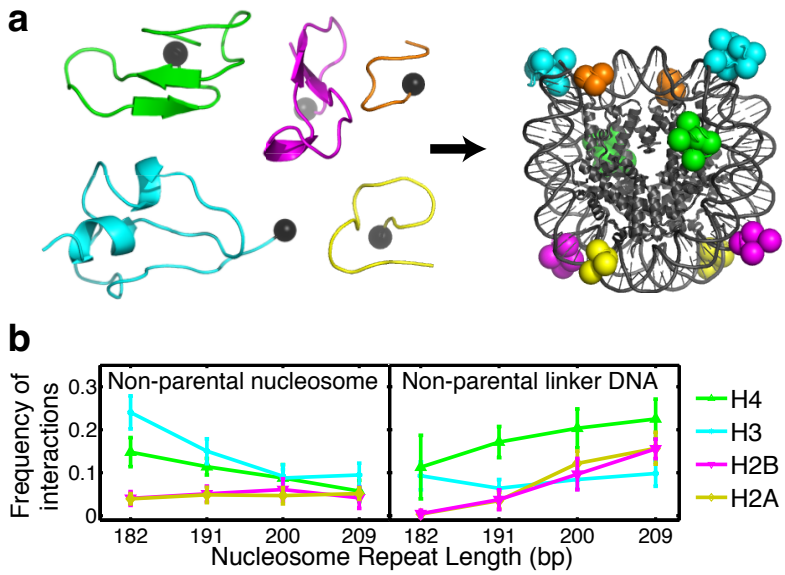

Figure S9. Modelling of histone tail folding and role of histone tails vs NRL. (a) A cartoon depicting the incorporation of the most populated folded histone tail conformation into the coarse-grained with histone tails in green (H4), cyan (H3), magenta (H2B), yellow (H2A), and orange (H2A C-tail). (b) Role of four different histone tails in mediating internucleosome interactions (i.e. the contacts between histone tails and non-parent nucleosomes or non-parent linker DNA) as a function of the nucleosome repeat length (147 bp of nucleosomal DNA plus the variable linker DNA length). In the 182 bp (35-bp linker-DNA length) arrays, the H3 and H4 tails spend more time mediating interactions with neighbouring nucleosomes, than with non-parental DNA linkers, for the 209-bp (62-bp linker-DNA length) arrays they engage more in interactions with non-parental DNA linkers.

Supplementary material bibliography

1 Arya, G. & Schlick, T. Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 16236-16241, doi:10.1073/pnas.0604817103 (2006).

2 Zhang, Q., Beard, D. A. & Schlick, T. Constructing irregular surfaces to enclose macromolecular complexes for mesoscale modeling using the discrete surface charge optimization (DISCO) algorithm. *Journal of computational chemistry* **24**, 2063-2074, doi:10.1002/jcc.10337 (2003).

3 Arya, G. & Schlick, T. A tale of tails: how histone tails mediate chromatin compaction in different salt and linker histone environments. *The journal of physical chemistry. A* **113**, 4045-4059, doi:10.1021/jp810375d (2009).

4 Arya, G., Zhang, Q. & Schlick, T. Flexible histone tails in a new mesoscopic oligonucleosome model. *Biophysical journal* **91**, 133-150, doi:10.1529/biophysj.106.083006 (2006).

5 Beard, D. A. & Schlick, T. Computational modeling predicts the structure and dynamics of chromatin fiber. *Structure* **9**, 105-114 (2001).

6 Beard, D. A. & Schlick, T. Modeling salt-mediated electrostatics of macromolecules: the discrete surface charge optimization algorithm and its application to the nucleosome. *Biopolymers* **58**, 106-115, doi:10.1002/1097-0282(200101)58:1<106::AID-BIP100>3.0.CO;2-# (2001).

7 Collepardo-Guevara, R. & Schlick, T. The effect of linker histone's nucleosome binding affinity on chromatin unfolding mechanisms. *Biophysical journal* **101**, 1670-1680, doi:10.1016/j.bpj.2011.07.044 (2011).

8 Collepardo-Guevara, R. & Schlick, T. Crucial role of dynamic linker histone binding and divalent ions for DNA accessibility and gene regulation revealed by mesoscale modeling of oligonucleosomes. *Nucleic acids research* **40**, 8803-8817, doi:10.1093/nar/gks600 (2012).

9 Collepardo-Guevara, R. & Schlick, T. Chromatin fiber polymorphism triggered by variations of DNA linker lengths. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 8061-8066, doi:10.1073/pnas.1315872111 (2014).

10 Perisic, O., Collepardo-Guevara, R. & Schlick, T. Modeling studies of chromatin fiber structure as a function of DNA linker length. *Journal of molecular biology* **403**, 777-802, doi:10.1016/j.jmb.2010.07.057 (2010).

11 Sun, J., Zhang, Q. & Schlick, T. Electrostatic mechanism of nucleosomal array folding revealed by computer simulation. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 8180-8185, doi:10.1073/pnas.0408867102 (2005).

12 Luque, A., Collepardo-Guevara, R., Grigoryev, S. & Schlick, T. Dynamic condensation of linker histone C-terminal domain regulates chromatin structure. *Nucleic acids research*, doi:10.1093/nar/gku491 (2014).

13 Stigter, D. Interactions of highly charged colloidal cylinders with applications to double-stranded. *Biopolymers* **16**, 1435-1448, doi:10.1002/bip.1977.360160705 (1977).

14 Allison, S. Brownian dynamics simulation of wormlike chains. Fluorescence depolarization and depolarized light scattering. *Macromolecules* **19**, 363-375 (1986).

15 Jian, H., Schlick, T. & Vologodskii, A. Internal motion of supercoiled DNA: brownian dynamics simulations of site juxtaposition. *Journal of molecular biology* **284**, 287-296, doi:10.1006/jmbi.1998.2170 (1998).

16 Jian, H., Vologodskii, A. & Schlick, T. A Combined Wormlike-Chain and Bead Model for Dynamic Simulations of Long Linear DNA. *Journal of Computational Physics* **136**, 168-179 (1997).

17 Siepmann, J. I. & Frenkel, D. Configurational Bias Monte-Carlo - a New Sampling Scheme for Flexible Chains. *Mol Phys* **75**, 59-70, doi:Doi 10.1080/00268979200100061 (1992).

18 Rosenbluth, M. N. & Rosenbluth, A. W. Monte-Carlo Calculation of the Average Extension of Molecular Chains. *Journal of Chemical Physics* **23**, 356-359 (1955).

19 Schuler, B., Lipman, E. A., Steinbach, P. J., Kumke, M. & Eaton, W. A. Polyproline and the "spectroscopic ruler" revisited with single-molecule fluorescence. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 2754-2759, doi:Doi 10.1073/Pnas.0408164102 (2005).

20 Papamokos, G. V. *et al.* Structural role of RKS motifs in chromatin interactions: a molecular dynamics study of HP1 bound to a variably modified histone tail. *Biophysical journal* **102**, 1926-1933, doi:10.1016/j.bpj.2012.03.030 (2012).

21 Grauffel, C., Stote, R. H. & Dejaegere, A. Force field parameters for the simulation of modified histone tails. *Journal of computational chemistry* **31**, 2434-2451, doi:10.1002/jcc.21536 (2010).

22 Camilloni, C., De Simone, A., Vranken, W. F. & Vendruscolo, M. Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry* **51**, 2224-2231, doi:10.1021/bi3001825 (2012).
23 Robinson, P. J. J. *et al.* 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. *Journal of molecular biology* **381**, 816-825, doi:Doi 10.1016/J.Jmb.2008.04.050 (2008).
24 Best, R. B. & Hummer, G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *The journal of physical chemistry. B* **113**, 9004-9015, doi:10.1021/jp901540t (2009).
25 Camilloni, C., Cavalli, A. & Vendruscolo, M. Assessment of the Use of NMR Chemical Shifts as Replica-Averaged Structural Restraints in Molecular Dynamics Simulations to Characterize the Dynamics of Proteins. *J. Phys. Chem. B* **117**, 1838-1843, doi:Doi 10.1021/Jp3106666 (2013).